

The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method

J.S. Spencer,^{1,2} N.S. Blunt,² and W.M.C. Foulkes²

¹*Department of Materials, Imperial College London, Exhibition Road, London, SW7 2AZ, U.K.*

²*Department of Physics, Imperial College London, Exhibition Road, London, SW7 2AZ, U.K.*

The recently proposed full configuration interaction quantum Monte Carlo method allows access to essentially exact ground-state energies of systems of interacting fermions substantially larger than previously tractable without knowledge of the nodal structure of the ground-state wave function. We investigate the nature of the sign problem in this method and how its severity depends on the system studied. We explain how cancellation of the positive and negative particles sampling the wave function ensures convergence to a stochastic representation of the many-fermion ground state and accounts for the characteristic population dynamics observed in simulations.

I. INTRODUCTION

One of the major goals of electronic structure methods is to produce accurate ground-state energies and properties of many-electron systems.¹ Quantum chemistry provides a hierarchy of *ab initio* methods² based upon Hartree-Fock, of which coupled-cluster singles and doubles with perturbative triples (CCSD(T)) is the most accurate applicable to medium-sized molecules.³ Quantum Monte Carlo methods such as Green's function Monte Carlo^{4,5} (GFMC), the closely related diffusion Monte Carlo⁶⁻⁸ (DMC), and auxiliary-field quantum Monte Carlo⁹ (AFQMC) produce accurate results via a stochastic sampling of the many-electron wave function, but none of these methods is exact: GFMC and DMC simulations of all but the smallest systems converge to the physically irrelevant many-boson ground state unless the fixed-node approximation is made;^{6,7} whilst AFQMC requires the phaseless approximation¹⁰ in order to avoid an exponential growth in noise except in certain special cases.¹¹

The full configuration interaction (FCI) method¹² casts the Schrödinger equation as a matrix eigenvalue problem, in which the requirement that the many-electron wave function be anti-symmetric with respect to exchange of electrons is imposed by working in a space of Slater determinants formed from a finite basis set of single-particle wave functions. The lowest eigenvalue and eigenvector of the FCI Hamiltonian matrix give the exact ground-state energy and wave function of the system, subject only to the error due to the finite basis set. Whilst the computational cost of FCI scales factorially with system size, it nevertheless represents the holy grail of electronic structure methods.

In 2009, Booth, Thom and Alavi¹³ introduced a new stochastic approach in which the nodal structure of the ground-state wave function emerges spontaneously by sampling the discrete space of Slater determinants. Their “full configuration interaction quantum Monte Carlo” (FCIQMC) method yields exact (i.e. FCI-quality) results whilst requiring a fraction of the memory of an FCI cal-

culation using the same basis. The memory required by FCIQMC simulations still scales factorially with system size, but the exponent appears to be substantially smaller than for FCI simulations. Moreover, unlike the iterative diagonalisation schemes required for FCI, the FCIQMC algorithm is readily parallelizable and can run efficiently on thousands of cores. Alavi and co-workers have used FCIQMC to reproduce essentially every molecular FCI calculation ever done and have obtained ground-state energies for systems with Hilbert spaces many orders of magnitude larger than the largest FCI calculations.¹³⁻¹⁵ We believe that the FCIQMC method will become increasingly important in the electronic structure community, especially if improved algorithms or substantially cheaper approximations can be found. Our motivation for exploring the behavior of the method is to provide insight into possible improvements.

An FCI calculation based on iterative diagonalization (using, for example, the Davidson method) requires the storage of at least two vectors, each of length equal to the size of the Hilbert space. An FCIQMC simulation using the same basis requires the storage of the labels of the determinants occupied by a population of stochastic walkers (which, following Anderson,⁶ we call “psi-particles” or *psips*) scattered over the same Hilbert space. In order for FCIQMC to be more efficient than FCI, the number of psips required must be a small fraction of the size of the Hilbert space. The fraction required is system dependent and provides a measure of how “hard” it is for FCIQMC to treat that system. The hardness is surprisingly difficult to predict. FCIQMC is wildly successful for some systems, such as the neon atom, where it requires only $\sim 0.01\%$ of the memory of a conventional FCI calculation.¹³ Even for the nitrogen molecule, a classic example of a strongly correlated system and a tough test for quantum chemical methods, FCIQMC used only a quarter of the memory of the equivalent FCI calculation.¹³ However, FCIQMC struggles to describe the methane molecule,¹³ for which Hartree-Fock is a very good approximation. We show in Sec. II that FCIQMC also struggles when applied to the Hubbard model and cannot treat systems larger than existing FCI methods

unless U is very small. What is it that makes a system difficult? Evidently the answer is more complicated than whether or not the system is strongly correlated.

The aim of this paper is to understand the FCIQMC algorithm better. Why are some systems more difficult to treat than others? What determines the characteristic population dynamics observed in FCIQMC simulations? How does the cancelation of positive and negative psips ensure convergence to the many-fermion ground state? How many psips are required to obtain correct results? What goes wrong if the population of psips is too small? We provide at least partial answers to all of these questions.

Section II reviews the FCIQMC method and applies it to the Hubbard model as an example. In Sec. III we discuss the nature of the sign problem in FCIQMC and explain the effect of canceling positive and negative psips. Section IV shows how our analysis of the sign problem also explains the characteristic population dynamics observed in FCIQMC simulations. Section V describes several special cases in which the sign problem can be removed entirely. We offer some concluding remarks in Sec. VI.

II. FCIQMC METHOD

We briefly summarize the FCIQMC method here, largely following the derivations given by Booth *et al.*,¹³ with attention paid to details relevant later in this paper.

Consider an orthonormal set of $2M$ single-particle spin-orbitals, $\{\phi_1, \phi_2, \dots, \phi_{2M}\}$. Previous FCIQMC work^{13–15} has used a basis of Hartree-Fock spin-orbitals, which is often a sensible choice but not required. One can construct an N -electron Slater determinant, D_i , by selecting any N spin-orbitals (assuming $N \leq 2M$):

$$D_i = D_{i_1, i_2, \dots, i_N} \quad (1)$$

$$= \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{i_1}(1) & \phi_{i_1}(2) & \cdots & \phi_{i_1}(N) \\ \phi_{i_2}(1) & \phi_{i_2}(2) & \cdots & \phi_{i_2}(N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{i_N}(1) & \phi_{i_N}(2) & \cdots & \phi_{i_N}(N) \end{vmatrix}. \quad (2)$$

To ensure that all such determinants are unique (up to a sign), it is convenient to require that $i_1 < i_2 < \dots < i_N$.

The ground-state wave function, Ψ_0 , may be defined variationally as the N -electron wave function that minimizes the energy expectation value,

$$E[\Psi] = \langle \Psi | \hat{H} | \Psi \rangle, \quad (3)$$

subject to the normalization constraint $\langle \Psi | \Psi \rangle = 1$. If we restrict the search for Ψ_0 to the subset of wave functions that can be expanded in the basis of ${}^{2M}C_N$ determinants,

$$\Psi_0 = \sum_i c_i D_i, \quad (4)$$

and impose the normalization constraint using a Lagrange multiplier, the optimal coefficients c_{0i} form the lowest eigenvector \mathbf{c}_0 of the matrix-eigenvalue problem

$$\sum_j \langle D_i | \hat{H} | D_j \rangle c_{0j} = \sum_j H_{ij} c_{0j} = E_0 c_{0i}, \quad (5)$$

where E_0 is both the Lagrange multiplier and the ground-state energy eigenvalue. This approach yields the FCI wave function for the finite basis set used.

One way to find the ground-state eigenvector \mathbf{c}_0 of the Hamiltonian matrix \mathbf{H} is to solve the imaginary-time Schrödinger equation:

$$\frac{d\mathbf{c}_i}{d\tau} = - \sum_j H_{ij} c_j. \quad (6)$$

The solution vector $\mathbf{c}(\tau)$ converges to \mathbf{c}_0 as $\tau \rightarrow \infty$ whenever the starting vector $\mathbf{c}(\tau=0)$ has a non-zero overlap with \mathbf{c}_0 ; indeed, this is also the driving principle behind other projector methods such as DMC and GFMC. The convergence is easy to demonstrate by considering the formal solution of the imaginary-time Schrödinger equation:

$$\mathbf{c}(\tau) = e^{-\mathbf{H}\tau} \mathbf{c}(0). \quad (7)$$

If the initial vector, $\mathbf{c}(0)$, is expanded in terms of the complete orthonormal set of the ${}^{2M}C_N$ eigenvectors, $\{\mathbf{c}_\alpha\}$, of the Hamiltonian matrix, \mathbf{H} ,

$$\mathbf{c}(0) = \sum_\alpha v_\alpha(0) \mathbf{c}_\alpha, \quad (8)$$

the solution of the imaginary-time Schrödinger equation can be written as

$$\mathbf{c}(\tau) = e^{-\mathbf{H}\tau} \sum_\alpha v_\alpha(0) \mathbf{c}_\alpha = \sum_\alpha v_\alpha(0) e^{-E_\alpha \tau} \mathbf{c}_\alpha. \quad (9)$$

The summation is dominated by the ground-state contribution in the limit $\tau \rightarrow \infty$ and thus

$$\mathbf{c}(\tau \rightarrow \infty) \approx v_0(0) e^{-E_0 \tau} \mathbf{c}_0. \quad (10)$$

The steady change in normalization due to the $e^{-E_0 \tau}$ factor is awkward but can be removed by choosing the zero of energy such that $E_0 = 0$. In practice, E_0 is not known until the end of the simulation, so we instead solve

$$\frac{d\mathbf{c}_i}{d\tau} = - \sum_j (H_{ij} - S \delta_{ij}) c_j, \quad (11)$$

where the energy shift, S , is adjusted slowly to keep the normalization more or less constant. The energy shift converges to the ground-state eigenvalue, E_0 , in the long-time limit. Note that previous work^{13–15} also subtracted the Hartree-Fock energy, E_{HF} , from the diagonal elements of the Hamiltonian, but this amounts merely to a redefinition of S . To make the subsequent notation as

simple as possible, we define a “transition matrix”, \mathbf{T} , via

$$\mathbf{T} = -(\mathbf{H} - S\mathbf{I}). \quad (12)$$

Note that the minimal eigenvalue of \mathbf{H} corresponds to the maximal eigenvalue of \mathbf{T} .

The FCIQMC algorithm proposed by Booth *et al.*¹³ may be summarized as follows. Consider a collection of markers distributed over the space of determinants, $\{D_i\}$. The markers do not move through the Hilbert space, so following Anderson⁶ we call them “psi particles” or *psips* instead of “walkers”. (It is easy to devise alternative FCIQMC algorithms in which the psips do move and behave much like walkers in DMC,⁸ but we use the approach of Booth, Thom and Alavi here.) Each psip has both a location, \mathbf{i} , and a “charge”, $q = \pm 1$, associated with it.

In one time step, $\Delta\tau$, we loop over the population of psips and allow each to “spawn” children (new psips) located on new determinants (which may be the same as the parent’s determinant) according to the following rules:

- The probability that a psip at \mathbf{i} successfully spawns a child at \mathbf{j} is $|T_{ji}|\Delta\tau$.
- If the spawning event is successful, the child psip has charge $q_{\text{child}} = \text{sign}(T_{ji})q_{\text{parent}}$, where q_{parent} is the charge of the parent.

At the end of the time step, pairs of psips of opposite charge on the same determinant cancel each other out (“annihilate”) and are removed from the simulation. For example, psips on determinant D_i with negative diagonal elements T_{ii} may generate children which immediately annihilate with the parent psips. Thus, although psips on different determinants may have different signs, all psips on any given determinant have the same sign at the end of every time step. Any psips which remain (whether originally “child” or “parent”) are merged and are allowed to spawn new psips in subsequent time steps.

The FCIQMC algorithm¹³ actually yields a stochastic sampling of the solution of the iterative equation

$$c_i(\tau + \Delta\tau) = \sum_j (\delta_{ij} + T_{ij}\Delta\tau)c_j(\tau), \quad (13)$$

which may be regarded as a finite-difference approximation to the imaginary-time Schrödinger equation

$$\frac{dc_i}{d\tau} = \sum_j T_{ij}c_j. \quad (14)$$

Assuming that $S = E_0$, as is the case once the simulation has converged on the ground state, the dominant eigenvectors of $\delta_{ij} + T_{ij}\Delta\tau$ and T_{ij} are identical if $\Delta\tau(E_{\text{max}} - E_0) < 2$, where E_{max} is the largest eigenvalue of the FCI Hamiltonian matrix \mathbf{H} . This means that there is no time-step error if $\Delta\tau$ is small enough. In order for there to be a finite time step, the Hamiltonian must be

bounded from both above and below. It also follows that the time step is limited if high-energy states are included in the basis; for example increasing the basis set in calculations of quantum chemical systems requires a decrease in the time step.

Rather than considering all possible spawning events, it is computationally efficient to allow a psip at D_i to attempt to spawn only onto its own determinant and a single connected determinant D_j per time step. (Connected determinants are linked by non-zero transition matrix elements T_{ji} .) The connected determinant is chosen stochastically according to some *generation probability* $p_{\text{gen}}(\mathbf{j} \leftarrow \mathbf{i})$, which must be non-zero whenever T_{ji} is non-zero and normalized such that $\sum_j p_{\text{gen}}(\mathbf{j} \leftarrow \mathbf{i}) = 1$. To maximise the probability of successful spawning, the generation probabilities should be as similar to $|T_{ji}|/\sum_j |T_{ji}|$ as possible; in practice, we go for speed and set the generation probabilities for spawning on all connected determinants equal. Once the candidate spawning location \mathbf{j} has been chosen, the spawning event is accepted with probability $|T_{ji}|\Delta\tau/p_{\text{gen}}(\mathbf{j} \leftarrow \mathbf{i})$.¹³ The effective size of the Hilbert space can be reduced by only selecting candidate spawning locations of some desired symmetry.

The population evolution of the psips over the course of an FCIQMC simulation has some commonly observed features, as illustrated in Fig. 1; we develop simple mathematical models to understand these results in Sections III and IV. The energy shift S is initially set to a value greater than (i.e., less negative than) the ground-state energy eigenvalue, ensuring that the largest eigenvalue of $\mathbf{T} = S\mathbf{I} - \mathbf{H}$ is positive and hence that total psip population grows exponentially. Once the population reaches a critical size the simulation spontaneously enters a “plateau” regime, where the total psip population remains stable and during which the ground-state wave function emerges. The height of the plateau relative to the size of the space of Slater determinants provides a measure for how hard a system is to solve using FCIQMC. After a system- and S -dependent waiting time, the simulation exits the plateau phase and the psip population starts to grow in an exponential fashion again, albeit at a slower rate than before. At this point the distribution of psips is a stochastic representation of the FCI ground-state wave function. Allowing the psip population to increase further serves only to reduce the statistical noise, so it is convenient to hold the population constant by varying the shift from this point in the simulation. The changes in shift must be carried out smoothly and slowly to avoid introducing a bias. Following Booth,¹³ we use the shift-update algorithm originally proposed for DMC simulations by Umrigar¹⁶:

$$S(\tau + A\Delta\tau) = S(\tau) + \frac{\xi}{A\Delta\tau} \ln \left(\frac{N_p(\tau + A\Delta\tau)}{N_p(\tau)} \right), \quad (15)$$

where A is the number of iterations between which the shift is updated, ξ is a damping parameter and $N_p(\tau)$ is the total psip population at time τ . In the simulations presented here we set $A = 20$ and $\xi = 0.1$. Once the psip

population has settled down, the expected value \bar{q}_i of the sum of the charges of the psips on D_i is proportional to the weight of D_i in the exact ground-state wave function:

$$\lim_{\tau \rightarrow \infty} \bar{q}_i(\tau) \propto c_{0i}. \quad (16)$$

Note that annihilation of psips of opposite charge does not affect \bar{q}_i .

There are two main ways of obtaining an estimate of the ground-state energy from an FCIQMC simulation. The *shift estimator* is the mean value of the shift S in the constant-population-mode simulation after the plateau. The *projected estimator* is obtained by noting that, for any single determinant D_0 with a non-zero ground-state component, the quantity

$$E = \lim_{\tau \rightarrow \infty} \frac{\langle D_0 | \hat{H} | \Psi(\tau) \rangle}{\langle D_0 | \Psi(\tau) \rangle} \quad (17)$$

$$= \frac{\sum_j H_{0j} \bar{q}_j}{\bar{q}_0} \quad (18)$$

tends to the ground-state energy as $\tau \rightarrow \infty$. Since the Hamiltonian operator only links determinants differing by two or fewer excitations, the number of terms included in the sum is limited. Note that it is important to average $q_j(\tau)$ and $q_0(\tau)$ separately; the projected estimator and its associated error can then be found by taking the ratio of the averages and using the covariance, respectively. Taking D_0 to be the determinant with the largest overlap with the exact ground-state wave function minimises the relative stochastic noise in the denominator of the above equation. Furthermore, such a determinant will typically have single and double excitations which also have significant contributions to the ground-state wave function, and hence determinants contributing to the numerator will also often have a significant population. The determinant with greatest overlap may not necessarily be known *a priori* (or even be clearly defined, as is the case in systems studied in Sec. V) but in practice the Hartree-Fock determinant is usually a good choice. The choice of D_0 can be changed during the course of a simulation if a determinant with a particularly large population is found.

We can illustrate the main features of the FCIQMC method by applying it to the Hubbard Hamiltonian¹⁷ for a d -dimensional cubic lattice,

$$\hat{H} = -t \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle, \sigma} \hat{c}_{\mathbf{r}, \sigma}^\dagger \hat{c}_{\mathbf{r}', \sigma} + U \sum_{\mathbf{r}} \hat{n}_{\mathbf{r}, \uparrow} \hat{n}_{\mathbf{r}, \downarrow}, \quad (19)$$

where $\hat{c}_{\mathbf{r}, \sigma}^\dagger$ ($\hat{c}_{\mathbf{r}, \sigma}$) creates (destroys) an electron of spin σ on site \mathbf{r} , the number operator $\hat{n}_{\mathbf{r}, \sigma} = \hat{c}_{\mathbf{r}, \sigma}^\dagger \hat{c}_{\mathbf{r}, \sigma}$ counts the spin- σ electrons on site \mathbf{r} , and the summation over $\langle \mathbf{r}, \mathbf{r}' \rangle$ includes all nearest-neighbor pairs of lattice sites. Assuming a system of M sites and $2M$ spin-orbitals subject to periodic boundary conditions, the Hubbard Hamiltonian

may also be written in reciprocal space,

$$\hat{H} = \sum_{\mathbf{k}, \sigma} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}, \sigma}^\dagger \hat{c}_{\mathbf{k}, \sigma} + \frac{U}{M} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} \hat{c}_{\mathbf{k}_1, \uparrow}^\dagger \hat{c}_{\mathbf{k}_2, \downarrow}^\dagger \hat{c}_{\mathbf{k}_3, \downarrow} \hat{c}_{\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3, \uparrow}, \quad (20)$$

where $\hat{c}_{\mathbf{k}, \sigma}^\dagger = \frac{1}{\sqrt{M}} \sum_{\mathbf{r}} \hat{c}_{\mathbf{r}, \sigma}^\dagger e^{i\mathbf{k} \cdot \mathbf{r}}$ is the creation operator for a Bloch state of wave vector \mathbf{k} , the sums are over the M wave vectors in the first Brillouin zone, $\epsilon_{\mathbf{k}} = -2t \sum_{i=1}^d \cos(k_i a)$ is the one-electron kinetic energy for wave vector $\mathbf{k} = (k_1, k_2, \dots, k_d)$, and a is the lattice parameter (which is set to unity from now on). The combination of wave vectors $\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3$ is assumed to have been reduced into the first Brillouin zone by addition of a reciprocal lattice vector if necessary.

Whilst the FCIQMC algorithm may be used in the real- and reciprocal-space representations, here, for illustrative purposes, we focus on the reciprocal-space formulation, as would be appropriate if U/t were small. As U/t increases and so electrons become increasingly more localised, the real-space formulation becomes more appropriate. Indeed, as we show in Sec. V, the real-space basis is substantially more favourable for non-zero values of U/t in the one-dimensional Hubbard model.

The imaginary-time evolution of the psip population and energy estimators during a \mathbf{k} -space FCIQMC simulation of the 2D 18-site half-filled Hubbard model at $U/t = 4$ (which is *not* small) is shown in Fig. 1. As might be expected, the \mathbf{k} -space FCIQMC method becomes less efficient at finding the ground state of this system as U/t increases — in fact, the psip population at the plateau increases approximately linearly with U/t for $U/t > 2$ (Fig. 2). When $U/t \geq 4$, one needs at least as many psips as there are determinants in the Hilbert space and the memory requirements are comparable to those of iterative diagonalization. As a result, unless U/t is small, \mathbf{k} -space FCIQMC is not able to treat half-filled Hubbard model systems substantially larger than those accessible to the FCI method.

Working in a space of Slater determinants has two main advantages. As the basis is anti-symmetric with respect to exchange of two spin-orbitals, there is no need to use the fixed-node approximation to prevent collapse to the bosonic ground state, as in DMC and GFMC. Although FCIQMC still has a sign problem (see Sec. III), the instability is not with respect to the a bosonic state and is often less severe. The second advantage is that the annihilation of psips with opposite charges proves surprisingly efficient in the discrete space of Slater determinants. Walker cancellation can in principle cure the sign problem in continuum DMC and GFMC simulations too,^{21–30} but such algorithms are substantially more complicated and less successful than the sign cancellation in FCIQMC.

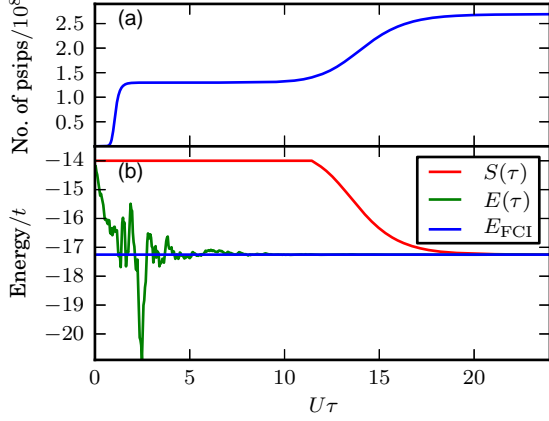


FIG. 1. Population dynamics and energy estimators for the 2D 18-site square lattice half-filled Hubbard model at $U = 4t$. Periodic boundary conditions are applied to a 45° tilted square cell with sides $3\sqrt{2}$. The ground-state wave function has momentum $\mathbf{k} = (0, 0)$ and the Hilbert space formed from all determinants of this momentum contains 1.3×10^8 states. Graph (a) shows how the psip population evolves with simulation time. The psip population initially grows exponentially before reaching a plateau, during which the distribution of psips settles to model the FCI wave function. The psip population then begins to grow exponentially again, reinforcing the wave function signal, at which point the shift is allowed to vary to stabilize the number of psips. Graph (b) shows the shift and projected energy estimators as a function of simulation time. The FCI energy is from Ref. 18. FCIQMC requires roughly the same number of psips as the size of the Hilbert space to converge on the ground-state wave function in this case.

III. THE ORIGIN OF THE SIGN PROBLEM IN FCIQMC

Booth *et al.*¹³ showed that annihilation is crucial in FCIQMC; without it the simulation never converges to the FCI ground state. In this section we attempt to explain why annihilation is required and how it helps the ground state to emerge.

First, consider the effect of removing annihilation from the simulation procedure. This means that psips of both charges are permitted to reside on and propagate from the same determinant at the same time. We shall use n_i^+ and n_i^- to represent the populations of positive and negative psips on determinant D_i . It is convenient to write the transition matrix \mathbf{T} as $\mathbf{T}^+ - \mathbf{T}^-$, where \mathbf{T}^+ contains the positive transition matrix elements, $T_{ij}^+ = \max(T_{ij}, 0)$, and \mathbf{T}^- contains the absolute values of the negative elements, $T_{ij}^- = \max(-T_{ij}, 0)$. All elements of \mathbf{T}^+ and \mathbf{T}^- are therefore non-negative. The populations of positive and negative psips evolve according to the

coupled differential equations:

$$\begin{aligned} \frac{dn_i^+}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^+ + T_{ij}^- n_j^- \right), \\ \frac{dn_i^-}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^- + T_{ij}^- n_j^+ \right). \end{aligned} \quad (21)$$

These can be decoupled by adding and subtracting:

$$\begin{aligned} \frac{d(n_i^+ + n_i^-)}{d\tau} &= \sum_j \left(T_{ij}^+ + T_{ij}^- \right) \left(n_j^+ + n_j^- \right), \\ \frac{d(n_i^+ - n_i^-)}{d\tau} &= \sum_j \left(T_{ij}^+ - T_{ij}^- \right) \left(n_j^+ - n_j^- \right). \end{aligned} \quad (22)$$

As $\tau \rightarrow \infty$, $\mathbf{n}^+ + \mathbf{n}^-$ tends to the eigenvector corresponding to the largest eigenvalue of $\mathbf{T}^+ + \mathbf{T}^-$, whilst $\mathbf{n}^+ - \mathbf{n}^-$ tends to the eigenvector corresponding to the largest eigenvalue of $\mathbf{T}^+ - \mathbf{T}^-$. We wish to find the latter state. However, as explained below, the largest eigenvalue of $\mathbf{T}^+ + \mathbf{T}^-$ is *always larger* than that of $\mathbf{T}^+ - \mathbf{T}^-$. Thus, the signal $\mathbf{n}^+ - \mathbf{n}^-$ *always decays* relative to $\mathbf{n}^+ + \mathbf{n}^-$. Fig. 3 shows that performing an FCIQMC simulation without annihilation does indeed give this undesired state. We note that a similar analysis has been performed previously for the world-line Quantum Monte Carlo method^{31,32}.

One can prove that the largest eigenvalue of $\mathbf{T}^+ + \mathbf{T}^-$ is always greater than or equal to the largest eigenvalue of $\mathbf{T}^+ - \mathbf{T}^-$ as follows. Suppose that \mathbf{c}_0 is the eigenvector corresponding to the largest eigenvalue, $\lambda_{\text{diff}}^{\text{max}}$, of $\mathbf{T}^+ - \mathbf{T}^-$. If \mathbf{c}_0 is normalized and the Hamiltonian matrix is real, all components of \mathbf{c}_0 may also be chosen real and so

$$\lambda_{\text{diff}}^{\text{max}} = \sum_{i,j} c_{0i} (T_{ij}^+ - T_{ij}^-) c_{0j}. \quad (23)$$

Now consider the vector $|\mathbf{c}_0|$ with components $|c_{0i}|$. This vector is also normalized and may be used as a trial state in the (upside-down) variational principle for the largest eigenvalue, $\lambda_{\text{sum}}^{\text{max}}$, of $\mathbf{T}^+ + \mathbf{T}^-$:

$$\lambda_{\text{sum}}^{\text{max}} \geq \sum_{i,j} |c_{0i}| (T_{ij}^+ + T_{ij}^-) |c_{0j}|. \quad (24)$$

The right-hand side of Eq. 24 is manifestly greater than or equal to the right-hand side of Eq. 23, so $\lambda_{\text{sum}}^{\text{max}} \geq \lambda_{\text{diff}}^{\text{max}}$.

The discrete annihilation process is difficult to model exactly³³ in a differential equation. We therefore consider a simpler annihilation process in which pairs of psips of opposite sign on the same determinant annihilate each other at a constant rate 2κ , where κ is a small positive constant and the factor of 2 is introduced solely for algebraic convenience. This leads to the differential equations

$$\begin{aligned} \frac{dn_i^+}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^+ + T_{ij}^- n_j^- \right) - 2\kappa n_i^+ n_i^-, \\ \frac{dn_i^-}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^- + T_{ij}^- n_j^+ \right) - 2\kappa n_i^+ n_i^-. \end{aligned} \quad (25)$$

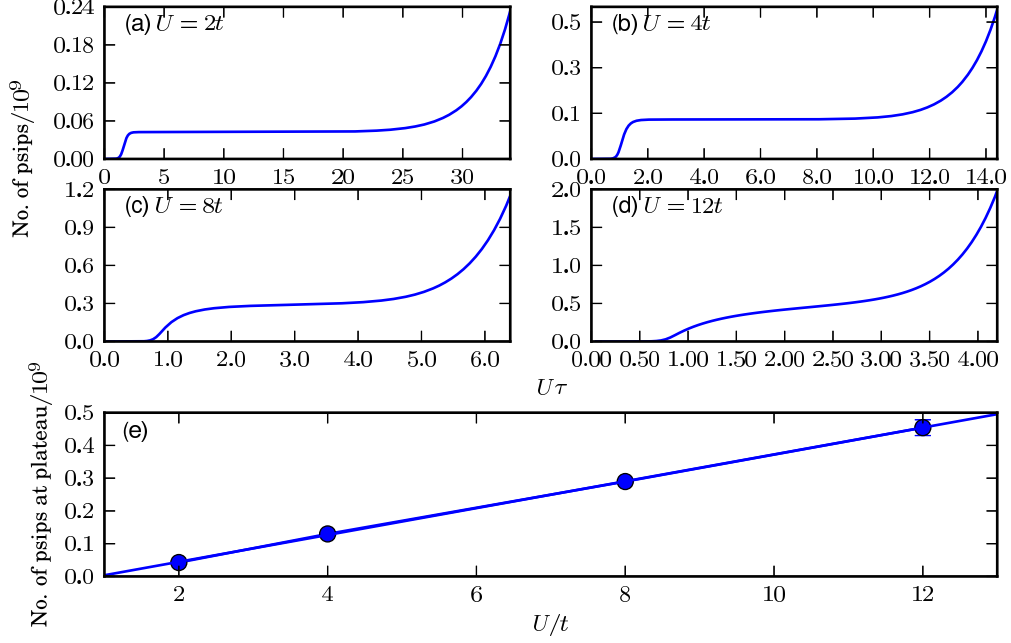


FIG. 2. Graphs (a)–(d) show the psip population dynamics for the 2D 18-site square lattice half-filled Hubbard model at $\mathbf{k} = (0,0)$ for various values of U/t . Periodic boundary conditions are applied to a 45° tilted square cell with sides $3\sqrt{2}$. The data for $U = 4t$ in (b) is identical to that in Fig. 1. Graph (e) shows that the number of psips at the plateau increases linearly with U/t when $U/t \geq 2$. The standard errors in the numbers of psips during the plateau phases were obtained using a blocking analysis¹⁹ and the straight line was fitted using the method of least squares as implemented in Ref. 20. Unless shown, the error bars are smaller than the markers.

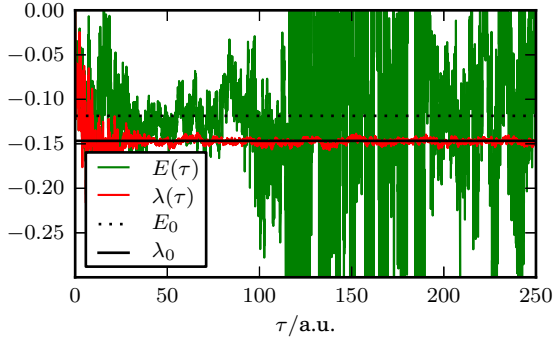


FIG. 3. The impact of removing annihilation from the FCIQMC algorithm applied to a 100×100 randomly-generated real symmetric Hamiltonian matrix \mathbf{H} . E_0 is the exact ground-state eigenvalue and $E(\tau)$ the projected estimator for \mathbf{H} . λ_0 and $\lambda(\tau)$ are the analogous quantities for the matrix $\mathbf{H}^+ + \mathbf{H}^-$, where $\mathbf{H} = \mathbf{H}^+ - \mathbf{H}^-$ and all elements of \mathbf{H}^+ and \mathbf{H}^- are negative (i.e., all elements of \mathbf{T}^+ and \mathbf{T}^- are positive). The simulation converges to the lowest eigenvector of $\mathbf{H}^+ + \mathbf{H}^-$, which does not correspond to an eigenvector of the Hamiltonian matrix.

The decoupled ordinary differential equations (ODEs) in

Eq. 22 thus become

$$\begin{aligned} \frac{d(n_i^+ + n_i^-)}{d\tau} &= \sum_j (T_{ij}^+ + T_{ij}^-) (n_j^+ + n_j^-) - 4\kappa n_i^+ n_i^-, \\ \frac{d(n_i^+ - n_i^-)}{d\tau} &= \sum_j (T_{ij}^+ - T_{ij}^-) (n_j^+ - n_j^-). \end{aligned} \quad (26)$$

It is clear how annihilation enables the FCIQMC method to converge upon the ground state of the Hamiltonian matrix: as the total psip population, $\sum_i (n_i^+ + n_i^-)$, rises, the rate of annihilation events rises quadratically. This causes the rate of growth of the $\mathbf{n}^+ + \mathbf{n}^-$ signal to decrease until spawning of new psips is balanced by annihilation. The growth of $\mathbf{n}^+ - \mathbf{n}^-$, the desired solution, is not affected and so this signal eventually emerges. As shown in Sec. IV, the rates of growth of the two signals actually become the same but due to annihilation all psips on the same determinant have the same sign.

To summarize, the sign problem in FCIQMC originates from the in-phase combination of positive and negative psips, which grows at a rate determined by the largest eigenvalue of $\mathbf{T}^+ + \mathbf{T}^-$. This eigenvalue is greater than the largest eigenvalue of $\mathbf{T}^+ - \mathbf{T}^-$, which determines the growth rate of the physical ground state of the system. The severity of the sign problem depends upon the dif-

ference between the largest eigenvalue of $\mathbf{T}^+ + \mathbf{T}^-$ and the largest eigenvalue of $\mathbf{T}^+ - \mathbf{T}^-$ — and thus on the prevalence of negative off-diagonal elements of \mathbf{T} — and upon the concentration of psips required to achieve a sufficient rate of annihilation for the growth of the in-phase signal to be suppressed. A difficult sign problem does not necessarily imply a strongly correlated system.

Eq. 21 models the time dependence of the populations of positive and negative psips in a simulation in which *all* annihilation is forbidden. In a real FCIQMC simulation, however, the annihilation rate per psip does not tend to zero as the psip density tends to zero. The probability of encountering and annihilating an unrelated psip vanishes, but a parent may still spawn a child of the opposite sign on its own determinant, in which case the parent and child annihilate each other. The low-density limit of the standard FCIQMC algorithm is better modeled by defining \mathbf{T}^+ and \mathbf{T}^- in a different way: \mathbf{T}^+ now contains *all* diagonal elements of \mathbf{T} , regardless of sign, and all positive off-diagonal matrix elements; whilst \mathbf{T}^- is zero on the diagonal but contains the absolute values of the negative off-diagonal elements. This redefinition corresponds to a change in viewpoint: instead of allowing parents to spawn children of the opposite sign on their own determinant, with subsequent annihilation, the negative diagonal elements of \mathbf{T}^+ remove psips from the simulation in one step, introducing an exponential decay of the psip population on determinants for which $T_{ii}^+ < 0$. The psip densities n_i^+ and n_i^- remain positive at all times and the above analysis is unchanged, but the severity of the sign problem, as measured by the difference between the largest eigenvalues of $\mathbf{T}^+ + \mathbf{T}^-$ and $\mathbf{T}^+ - \mathbf{T}^-$, is smaller than the above analysis suggests.

From now on it will be assumed that \mathbf{T}^+ and \mathbf{T}^- are defined as explained in the previous paragraph, and hence that all diagonal elements of \mathbf{T}^- are zero.

IV. POPULATION DYNAMICS

Separating the psip population into positive and negative contributions also explains the population dynamics observed during an FCIQMC simulation. Let $\mathbf{V} = \mathbf{T}^+ + \mathbf{T}^-$, $\mathbf{p} = \mathbf{n}^+ + \mathbf{n}^-$ and $\mathbf{n} = \mathbf{n}^+ - \mathbf{n}^-$. The decoupled ODEs in Eq. 26 can thus be written as

$$\begin{aligned} \frac{dp_i}{d\tau} &= \sum_j V_{ij}p_j - \kappa(p_i^2 - n_i^2) \\ \frac{dn_i}{d\tau} &= \sum_j T_{ij}n_j. \end{aligned} \quad (27)$$

In the long-time limit, $\mathbf{n}(\tau)$ becomes dominated by its ground-state component, which grows at a rate determined by the largest eigenvalue, $T_{\max} = S - E_0$, of the transition matrix \mathbf{T} :

$$\mathbf{n}(\tau) \approx \alpha e^{T_{\max}\tau} \mathbf{c}_0, \quad (28)$$

where $\alpha = \mathbf{c}_0 \cdot \mathbf{n}(0)$. Hence the evolution equation for \mathbf{p} becomes

$$\frac{dp_i}{d\tau} = \sum_j V_{ij}p_j - \kappa p_i^2 + \kappa \alpha^2 e^{2T_{\max}\tau} c_{0i}^2. \quad (29)$$

Eq. 29 is difficult to solve exactly, but the population dynamics it embodies can be illustrated using a simple one-component analogue:

$$\frac{dp}{d\tau} = V_{\max}p - \kappa p^2 + \kappa n^2, \quad (30)$$

where $p(\tau)$ is the total psip population at time τ , V_{\max} is larger than T_{\max} , and $n(\tau) = n_0 e^{T_{\max}\tau}$. It is common to start an FCIQMC simulation with a population of positive psips only, in which case the initial condition is $p(0) = n(0) = n_0$. As the initial psip population and annihilation rate are small, $p(\tau)$ grows exponentially at the start of the simulation: $p(\tau) = n_0 e^{V_{\max}\tau}$. The exponential growth ceases when $p \approx V_{\max}/\kappa$, at which point the κp^2 annihilation term (which is larger than the κn^2 term because $V_{\max} > T_{\max}$) counteracts it. The psip population then enters a plateau phase, remaining roughly constant until the $n(\tau)$ signal, which has been steadily growing like $n_0 e^{T_{\max}\tau}$, begins to dominate. From then on the population rises exponentially again, although now at a rate determined by T_{\max} . The ground-state wave function thus spontaneously emerges from the simulation.

The plateau begins to appear at a time τ_1 determined by the equation $p(\tau_1) \approx n_0 e^{V_{\max}\tau_1} \approx V_{\max}/\kappa$. Hence

$$\tau_1 \approx \frac{\ln(V_{\max}/\kappa n_0)}{V_{\max}}. \quad (31)$$

More precisely, solving Eq. 30 with the κn^2 term omitted and assuming that $\kappa n_0/V_{\max} \ll 1$ (which must be the case if the annihilation rate is negligible at the beginning of the simulation) shows that $p(\tau)$ reaches 95% of its plateau value at time

$$\tau_{95\%} = \frac{\ln(19(V_{\max}/\kappa n_0 - 1))}{V_{\max}}. \quad (32)$$

The end of the plateau occurs at a time τ_2 determined by the equation $n(\tau_2) = n_0 e^{T_{\max}\tau_2} \approx V_{\max}/\kappa$, and hence

$$\tau_2 \approx \frac{\ln(V_{\max}/\kappa n_0)}{T_{\max}}. \quad (33)$$

Combining Eqs. 31 and 33 yields

$$\frac{\tau_2}{\tau_1} \approx \frac{V_{\max}}{T_{\max}}. \quad (34)$$

The energy shift S appears on the diagonal of $\mathbf{T} = \mathbf{S}\mathbf{I} - \mathbf{H}$ and is incorporated into the diagonal elements of \mathbf{T}^+ ; it therefore affects $\mathbf{T} = \mathbf{T}^+ - \mathbf{T}^-$ and $\mathbf{V} = \mathbf{T}^+ + \mathbf{T}^-$ equally. This dependence can be made explicit by writing $T_{\max} = S + T_0$ and $V_{\max} = S + V_0$, where T_0 and $V_0 (\geq T_0)$

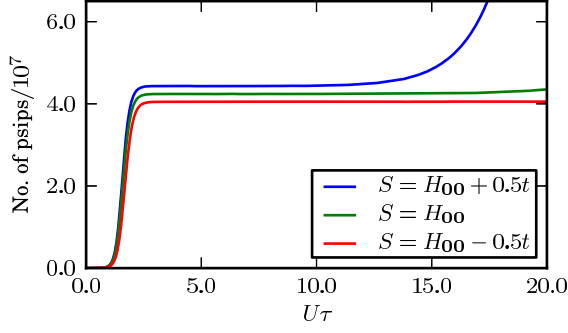


FIG. 4. Effect of the initial value of the shift on the plateau height in an FCIQMC simulation of the 2D 18-site square lattice half-filled Hubbard model at $U = 2t$ and $\mathbf{k} = (0, 0)$. Periodic boundary conditions are applied to a 45° tilted cell with sides $3\sqrt{2}$. The reference determinant, D_0 , is formed from occupying the 18 lowest-energy Bloch functions. Setting the shift slightly below $H_{00} = \langle D_0 | \hat{H} | D_0 \rangle$ reduces the plateau height but increases the time spent in the plateau region.

are the largest eigenvalues of \mathbf{T} and \mathbf{V} when $S = 0$. This leads to the equation

$$\frac{\tau_2}{\tau_1} \approx \frac{S + V_0}{S + T_0} = \frac{S + V_0}{S - E_0}. \quad (35)$$

In most FCIQMC simulations the shift is initially held at a fixed value, typically the Hartree-Fock energy, until the desired psip population is reached. If S is reduced and allowed to approach E_0 from above, the simulation never exits the plateau. The duration of the plateau can be shortened by moving the initial value of S further above the ground-state energy E_0 , but only at the cost of increasing the psip population $(S + V_0)/\kappa$ at the plateau. The effect of modifying the initial shift in an FCIQMC simulation is shown in Fig. 4.

The ODE in Eq. 30 is an example of a Riccati differential equation³⁴ and can be transformed from a quadratic first-order ODE into a linear second-order ODE using the substitution

$$p(\tau) = \frac{1}{\kappa u} \frac{du}{d\tau}. \quad (36)$$

The solution of the resultant second-order ODE is

$$u(\tau) = c_1 \cdot {}_0F_1 \left(; 1 - \frac{V_{\max}}{2T_{\max}}; z \right) + c_2 z^{V_{\max}/2T_{\max}} \cdot {}_0F_1 \left(; 1 + \frac{V_{\max}}{2T_{\max}}; z \right), \quad (37)$$

where

$$z = \frac{\kappa^2 p^2(0) e^{2T_{\max}\tau}}{4T_{\max}^2}, \quad (38)$$

c_1 and c_2 are constants of integration, ${}_0F_1$ is a confluent hypergeometric limit function,³⁵ and we have assumed

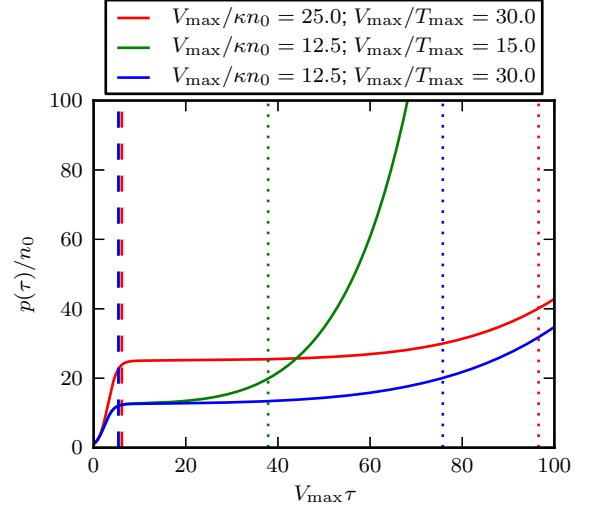


FIG. 5. Model population dynamics in an FCIQMC simulation. The psip population evolves according to Eq. 30 and shows the key features of the population dynamics in real FCIQMC calculations: an initial exponential growth phase followed by a plateau followed by a second slower exponential growth phase. As expected, the psip population reaches 95% of its plateau value when $V_{\max}\tau_{95\%} \approx \ln(19(V_{\max}/\kappa n_0 - 1))$, the height of the plateau is given by $p/n_0 = V_{\max}/\kappa n_0$, and the plateau ends when $V_{\max}\tau_2 \approx \ln(V_{\max}/\kappa n_0)V_{\max}/T_{\max}$. The vertical dashed (dotted) lines show $V_{\max}\tau_{95\%}$ ($V_{\max}\tau_2$) for each set of parameters. The hypergeometric function, ${}_0F_1$, was calculated using Ref. 20.

that $V_{\max}/(2T_{\max})$ is not an integer. The normalization of $u(\tau)$ drops out of $p(\tau) = (\kappa u)^{-1} du/d\tau$, leaving one arbitrary constant to be fixed by the initial conditions. Fig. 5 shows the resulting population dynamics for three different sets of parameters, imposing the boundary condition $p(0) = n_0$ in all cases. Whilst this is an extremely simple model, it captures all of the features of the population dynamics seen in actual FCIQMC calculations.

Let us now return to Eq. 29 and consider what happens beyond the end of the plateau, where the psip population begins to rise again until the energy shift is adjusted and a steady state with $dp_i/dt = 0$ is attained. Since the linear \mathbf{V} term is negligible in comparison with the quadratic terms in this regime, Eq. 29 becomes

$$0 = \frac{dp_i}{dt} \approx -\kappa p_i^2 + \kappa n_i^2, \quad (39)$$

implying that $p_i \approx \pm n_i$ and hence that $n_i^- \approx 0$ or $n_i^+ \approx 0$. Determinants on which the ground-state amplitude c_{0i} is positive are occupied only by positive psips and determinants on which c_{0i} is negative are occupied only by negative psips. The signed psip density $n_i = n_i^+ - n_i^-$ is proportional to c_{0i} and the unsigned density $p_i = n_i^+ + n_i^-$ to $|c_{0i}|$.

We can also use Eq. 29 to understand why the plateau psip populations plotted in Fig. 2 are proportional to the

Hubbard U . Assuming that the growth of the in-phase signal, \mathbf{p} , is fast enough to allow the plateau to emerge before the ground-state signal becomes significant, the plateau occurs when

$$\sum_{\mathbf{ij}} V_{\mathbf{ij}} p_{\mathbf{j}} \approx \kappa \sum_{\mathbf{i}} p_{\mathbf{i}}^2. \quad (40)$$

If U/t is large, the kinetic energy contributions to the Hamiltonian matrix (and hence \mathbf{T} and \mathbf{V}) can be neglected and changing U simply scales these matrices. Defining $V_{\mathbf{ij}} = UV'_{\mathbf{ij}}$ and $p_{\mathbf{i}} = Up'_{\mathbf{i}}$ leads to the following condition for the emergence of the plateau:

$$\sum_{\mathbf{ij}} V'_{\mathbf{ij}} p'_{\mathbf{j}} \approx \kappa \sum_{\mathbf{i}} p'^2_{\mathbf{i}}, \quad (41)$$

where \mathbf{V}' and \mathbf{p}' contain no dependence upon U . The total psip population at the plateau, $\sum_{\mathbf{i}} p_{\mathbf{i}} = U \sum_{\mathbf{i}} p'_{\mathbf{i}}$, therefore scales linearly with U .

V. SIGN-PROBLEM-FREE SYSTEMS

If there exists a similarity transform that maps \mathbf{T} into \mathbf{V} , and hence makes every off-diagonal element of \mathbf{T} positive and every off-diagonal element of $\mathbf{H} = \mathbf{S}\mathbf{I} - \mathbf{T}$ negative, then \mathbf{V} and \mathbf{T} have identical eigenvalues. Even without annihilation, the in-phase and out-of-phase signals, \mathbf{p} and \mathbf{n} , grow at the same rate and there is no difficulty extracting the ground-state (out-of-phase) signal $\mathbf{n} = \mathbf{n}^+ - \mathbf{n}^-$. Such systems are sign-problem free and FCIQMC simulations of them yield correct ground-state energies with arbitrarily small psip populations, although no plateau phase is seen. Increasing the psip population serves only to reduce the stochastic error. Indeed, by formulating the problem in the transformed basis, we can carry out an FCIQMC simulation in which no negative psips need ever appear.

A particularly simple type of similarity transformation does no more than change the signs of some of the basis determinants. No choice of signs is sufficient to render all off-diagonal elements of $T_{\mathbf{ij}}$ positive in most cases, but there are a few interesting model systems in which simple sign-changing transformations are effective. In some models, for example, all psips spawned on any given determinant $D_{\mathbf{i}}$ have the same sign regardless of the location of their parent, so that positive and negative psips never mix. A simple sign-changing transformation that multiplies every basis determinant by the sign of the psips that occupy it then makes all off-diagonal elements of \mathbf{T} positive.

The antiferromagnetic Heisenberg model defined by the Hamiltonian

$$\hat{H} = J \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \hat{\mathbf{S}}_{\mathbf{r}} \cdot \hat{\mathbf{S}}_{\mathbf{r}'}, \quad (42)$$

where $J > 0$, $\hat{\mathbf{S}}_{\mathbf{r}}$ is the vector spin operator on lattice site \mathbf{r} , and the sum runs over nearest neighbors, is such a system if the lattice is bipartite. (Note that the basis states

in this example are spin configurations rather than Slater determinants.) Indeed, the sign structure of the ground-state wave function of the bipartite Heisenberg model has long been known.³⁶ Every non-zero off-diagonal matrix element of the Heisenberg Hamiltonian flips a neighboring pair of spins from down-up to up-down or vice-versa. The Hamiltonian matrix element is positive if $J > 0$, implying that the corresponding matrix element of \mathbf{T} is negative, so the signs of children produced by off-diagonal spawning events always differ from the signs of their parents. Initially it appears as if this ought to cause a serious sign problem. Consider, however, a bipartite system consisting of two inter-penetrating sub-lattices, arranged such that the pairs of spins flipped by off-diagonal elements of \hat{H} are always on different sub-lattices. The action of any off-diagonal element of \hat{H} increases the total value of S_z on one sub-lattice by 1 and decreases the total value of S_z on the other sub-lattice by 1. This allows all spin configurations to be assigned to one of two classes, A or B. Starting from the classical Néel state, which is arbitrarily assigned to class A, all states that differ from the Néel state by an even number of applications of \hat{H} are also assigned to class A, while states that differ by an odd number of applications of \hat{H} are assigned to class B. If an FCIQMC simulation is initialized by placing a population of positive psips on the Néel state, all psips created on configurations in class A will be positive while all psips created on configurations in class B will be negative. Psips of different signs will never mix, no cancellation is required, and there is no sign problem. In fact, by applying a simple unitary transformation in which the sign of every configuration in class B is flipped, all off-diagonal matrix elements of \mathbf{H} can be made negative.³⁷

By way of an example, the Heisenberg model for a two-dimensional 6×6 square lattice with periodic boundary conditions has a Hilbert space of 9.08×10^9 configurations. Using just 1.8×10^6 psips we find the $E/N = -0.67886(2)J$, which is in excellent agreement with the value $E/N = -0.678871(8)J$ obtained by Runge using Green's function Monte Carlo.³⁸ Note that, unlike Runge, we did not use importance sampling. FCIQMC simulations of Heisenberg models defined on other (non-bipartite or "frustrated") lattices display the same behavior as for fermionic systems (Fig. 6); in particular, the population plateau is a universal feature unless the system is sign-problem free.

The Hubbard Hamiltonian, Eq. 19, is closely related to the Heisenberg Hamiltonian and is also sign-problem-free, although only in one dimension. Consider a 1D Hubbard lattice with N_s sites and N_{\uparrow} (N_{\downarrow}) spin-up (spin-down) electrons, where the system is not necessarily half-filled and periodic boundary conditions are applied. If N_s is even and N_{\uparrow} and N_{\downarrow} are both odd or N_s is odd and N_{\uparrow} and N_{\downarrow} are both even, then there exists an analogous transformation to that described above for the Heisenberg model. (This is most easily seen if the orbitals are ordered first by spin and then by their position in the lattice.) The ground-state energies of large 1D Hubbard

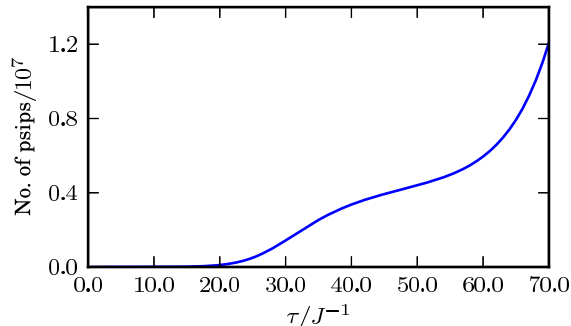


FIG. 6. The population dynamics of an FCIQMC simulation of the antiferromagnetic Heisenberg model for a 5×5 triangular lattice. No periodic boundary conditions were imposed. As with fermionic systems, frustrated Heisenberg models display a population plateau.

models with many different choices of N_{\uparrow} and N_{\downarrow} can therefore be found using the FCIQMC method with very small numbers of psips.

This brings us on to an important point: the sign problem in FCIQMC is not constant for a given system but is dependent upon the choice of basis. Consider two Hamiltonian matrices, \mathbf{H}_1 and \mathbf{H}_2 , which describe the same system but have been constructed using different basis sets, one of which is obtained from the other by a transformation matrix \mathbf{S} , so that $\mathbf{H}_1 = \mathbf{S}^{-1}\mathbf{H}_2\mathbf{S}$. The corresponding transition matrices \mathbf{T}_1 and \mathbf{T}_2 are related in the same way. However, when we partition the two transition matrices into \mathbf{T}^+ and \mathbf{T}^- , the partitions are not necessarily related by the same unitary transformation. In other words, although $\mathbf{T}_1 = \mathbf{S}^{-1}\mathbf{T}_2\mathbf{S}$, and hence $\mathbf{T}_1^+ - \mathbf{T}_1^- = \mathbf{S}^{-1}(\mathbf{T}_2^+ - \mathbf{T}_2^-)\mathbf{S}$, it is not in general the case that $\mathbf{T}_1^{\pm} = \mathbf{S}^{-1}\mathbf{T}_2^{\pm}\mathbf{S}$. Consequently, the plateau height (and thus the ease with which FCIQMC can be used to find the ground state) depends on the basis in which the Hamiltonian matrix is expressed.

An FCIQMC simulation of the 18-site half-filled 1D Hubbard model at $U = t$ exhibits no plateau when the Hamiltonian matrix is expressed in a basis of determinants of local orbitals (as in Eq. 19), since this particular problem is sign-problem free as explained above. Yet an FCIQMC simulation of exactly the same Hamiltonian expressed in a basis of determinants of Bloch functions (as in Eq. 20) exhibits a plateau at 6.9×10^6 psips. As expected, the sign problem increases the difficulty of the FCIQMC simulation in the Bloch representation. A short simulation using just 2.8×10^5 psips in the local orbital basis gave a ground-state energy of $-18.8423(3)t$, whereas 2.3×10^7 psips were required to obtain a ground-state energy of $-18.84248(8)t$ in the Bloch orbital basis. Furthermore, our implementation currently conserves crystal momentum when using Bloch orbitals but does not make use of symmetry when using local orbitals; the Hilbert space used in the calculation with local orbitals therefore contains 2.36×10^9 determinants,

whereas that with Bloch orbitals contains only 1.31×10^8 determinants.

VI. DISCUSSION

The above observations provide some degree of insight into the FCIQMC method and the factors that determine how hard it is to apply FCIQMC to any given physical system. We now briefly comment upon other topics related to FCIQMC before summarizing our work.

The initiator approximation to FCIQMC (i-FCIQMC), whereby spawning events onto previously unoccupied determinants are only allowed if the parent determinant has a psip population above a specified threshold, has been shown to dramatically reduce the number of psips required to obtain excellent results in many systems.¹⁵ In i-FCIQMC, the effective Hilbert space grows and changes dynamically as the simulation proceeds, including only determinants that have a significant psip population or lie close (in terms of applications of the Hamiltonian) to determinants with a significant psip population. As a result, psips are prevented from spawning in regions of low psip density and the annihilation rate is greatly increased for a given psip population. No plateau is observed in i-FCIQMC calculations^{14,15} because the growth of the in-phase combination is suppressed by annihilation even when the total psip population is low. The i-FCIQMC approximation becomes increasingly good as the number of psips is increased because more of the Hilbert space becomes accessible and the psip dynamics more closely resembles that of the true Hamiltonian.

In coupled cluster Monte Carlo³⁹ (CCMC), the excitation amplitudes of the coupled cluster wave function are stochastically sampled in manner analogous to the sampling of configuration amplitudes in FCIQMC. Although the CCMC algorithm is much more complicated than the FCIQMC algorithm, we believe that a similar analysis holds. The particles or “excips”⁴⁰ that sample the discrete excitor space in CCMC also have positive or negative signs and annihilation is crucial for the desired (and physically meaningful) solution to emerge. It is observed that CCMC calculations have a higher plateau than configuration interaction quantum Monte Carlo (CIQMC) calculations at the same truncation level.³⁹ It is likely that this is because the excips have to explore a larger effective Hilbert space than the psips in the equivalent CIQMC simulation, so more excips are required to make the annihilation rate high enough to suppress the in-phase signal.

In summary, we have shown that the sign problem in FCIQMC simulations is caused by the growth of an unphysical dominant solution of the coupled ODE’s that describe the time evolution of the densities of positive and negative psips. This unphysical solution grows faster than the ground-state solution we seek and masks the ground-state signal. A similar problem arises in DMC simulations of fermionic systems, where the dominant

solution is the boson ground state. FCIQMC differs from DMC, however, in that the unphysical solution in FCIQMC is not an eigenfunction of the Hamiltonian operator.

The annihilation of psips of opposite charge suppresses the growth of the unphysical dominant solution and allows FCIQMC simulations to converge on the true ground state, but works only when a minimum (and system-dependent) threshold in the psip population is exceeded. The combination of spawning and annihilation also leads to the characteristic population dynamics observed in FCIQMC calculations. The annihilation of psips may not be the most efficient way of suppressing the growth of the unphysical solution, and there remains a tantalising possibility that an alternative approach could lead to a superior algorithm applicable to much larger systems.

ACKNOWLEDGMENTS

JSS is grateful for discussions on coupled cluster Monte Carlo with A.J.W. Thom. This work made use of the Imperial College London High Performance Computing facilities.

- ¹P. Knowles, M. Schütz, and H.-J. Werner, in *Modern methods and algorithms of quantum chemistry*, Vol. 3, edited by J. Grotenndorst (John von Neumann Institute for Computing, 2000) pp. 97–179.
- ²C. Ochsenfeld, J. Kussmann, and D. S. Lambrecht, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and T. R. Cundari (John Wiley & Sons, Inc., 2007) pp. 1–82.
- ³R. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- ⁴M. H. Kalos, *Phys. Rev.* **128**, 1791 (1962).
- ⁵D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer-Verlag, 1979) pp. 145–194.
- ⁶J. B. Anderson, *J. Chem. Phys.* **63**, 1499 (1975).
- ⁷J. B. Anderson, *J. Chem. Phys.* **65**, 4121 (1976).
- ⁸W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **71**, 33 (2001).
- ⁹R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, *Phys. Rev. D* **24**, 2278 (1981).
- ¹⁰S. Zhang and H. Krakauer, *Phys. Rev. Lett.* **90**, 136401 (2003).
- ¹¹C. Wu and S.-C. Zhang, *Phys. Rev. B* **71**, 155115 (2005).
- ¹²P. J. Knowles and N. C. Handy, *Chem. Phys. Lett.* **111**, 315 (1984).
- ¹³G. H. Booth, A. J. W. Thom, and A. Alavi, *J. Chem. Phys.* **131**, 054106 (2009).
- ¹⁴G. H. Booth and A. Alavi, *J. Chem. Phys.* **132**, 174104 (2010).
- ¹⁵D. Cleland, G. H. Booth, and A. Alavi, *J. Chem. Phys.* **132**, 041103 (2010).
- ¹⁶C. J. Umrigar, M. P. Nightingale, and K. J. Runge, *J. Chem. Phys.* **99**, 2865 (1993).
- ¹⁷J. Hubbard, *Proc. R. Soc. London, Ser. A* **276**, pp. 238 (1963).
- ¹⁸F. Becca, A. Parola, and S. Sorella, *Phys. Rev. B* **61**, R16287 (2000).
- ¹⁹H. Flyvbjerg and H. G. Petersen, *J. Chem. Phys.* **91**, 461 (1989).
- ²⁰E. Jones, T. Oliphant, P. Peterson, *et al.*, “SciPy: Open source scientific tools for Python,” (2001–).
- ²¹D. M. Arnow, M. H. Kalos, M. A. Lee, and K. E. Schmidt, *J. Chem. Phys.* **77**, 5562 (1982).
- ²²J. Carlson and M. H. Kalos, *Phys. Rev. C* **32**, 1735 (1985).
- ²³D. F. Coker and R. O. Watts, *Mol. Phys.* **58**, 1113 (1986).
- ²⁴J. B. Anderson, C. A. Traynor, and B. M. Boghosian, *J. Chem. Phys.* **95**, 7418 (1991).
- ²⁵J. B. Anderson, *J. Chem. Phys.* **96**, 3702 (1992).
- ²⁶J. B. Anderson, C. A. Traynor, and B. M. Boghosian, *J. Chem. Phys.* **99**, 345 (1993).
- ²⁷S. Zhang and M. H. Kalos, *Phys. Rev. Lett.* **67**, 3074 (1991).
- ²⁸Z. Lui, S. Zhang, and M. H. Kalos, *Phys. Rev. E* **50**, 3220 (1994).
- ²⁹M. J. Kalos, *Phys. Rev. E* **53**, 5420 (1996).
- ³⁰M. J. Kalos and F. Pederiva, *Phys. Rev. Lett.* **85**, 3547 (2000).
- ³¹N. Hatano and M. Suzuki, *Phys. Lett. A* **163**, 246 (1992).
- ³²T. Nakamura, N. Hatano, and H. Nishimori, *J. Phys. Soc. Jpn.* **61**, 3494 (1992).
- ³³The effect can be modeled more accurately by using Heaviside step functions but the form shown here is more convenient.
- ³⁴E. L. Ince, *Ordinary differential equations*, 2nd ed. (Dover Publications, New York, 2001).
- ³⁵F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, eds., *NIST Handbook of Mathematical Functions* (Cambridge University Press, 2010).
- ³⁶W. Marshall, *Proc. R. Soc. A* **232**, 48 (1955).
- ³⁷E. Lieb and D. Mattis, *J. Math. Phys.* **3**, 749 (1962).
- ³⁸K. J. Runge, *Phys. Rev. B* **45**, 12292 (1992).
- ³⁹A. J. W. Thom, *Phys. Rev. Lett.* **105**, 263004 (2010).
- ⁴⁰A. J. W. Thom, private communication (2011).